Climate in a Box Workshop

Workflow Tool

NASA SIVO Code 610.3

September 22, 2010

Presenters:
March Andrews, Robert Burns, Megan Damon, John Evans, Shawn Freeman, Phil Hayes, Bruce Van Aartsen

Outline: Morning Session

- Introduction to the NASA Workflow Tool
- NASA Experiment Designer (NED)
- Break
- Running the GEOS5 Workflow
- Creating a Workflow – Part I: Overview
- Lunch

Outline: Afternoon Session

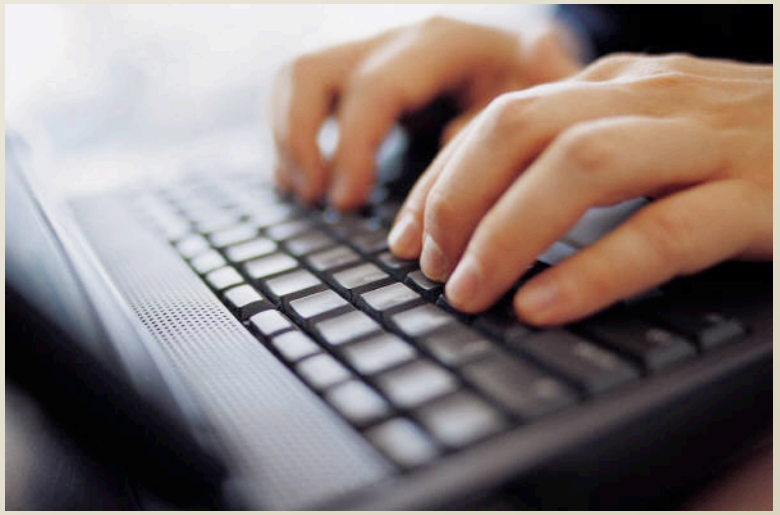
- Creating a Workflow – Part II: Hands-on Training
- Break
- Distributed Modeling System
- Workshop Summary

Introduction to the NASA Workflow Tool

What is a Workflow Tool?

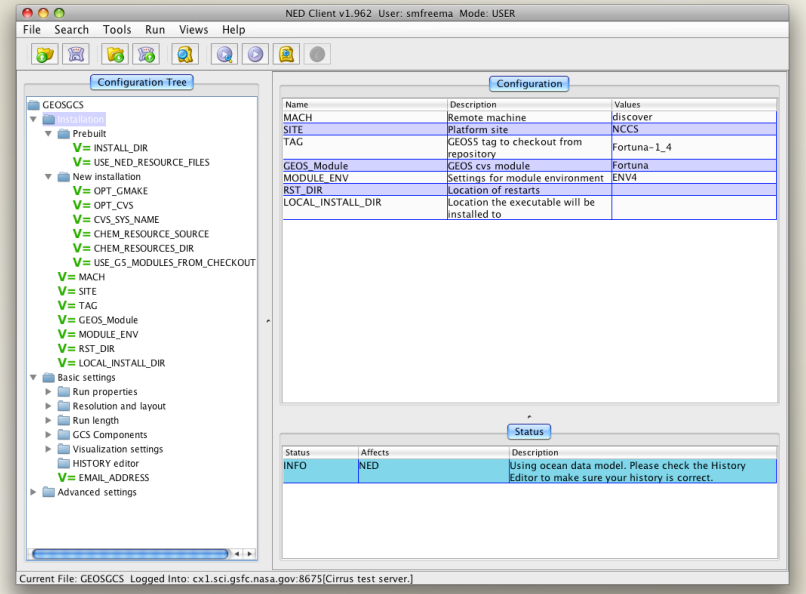
A system that supports automating a series of tasks such as setting up and running complex model experiments.

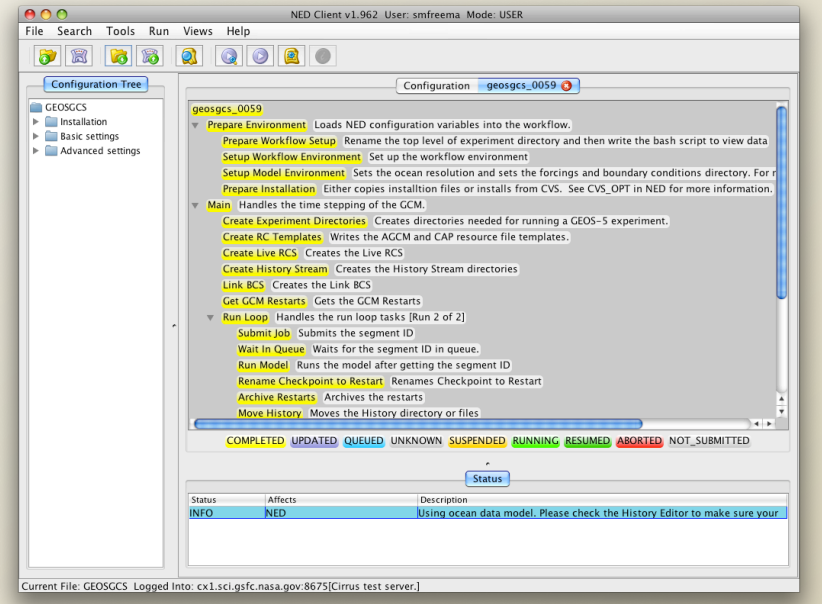
Enable users to concentrate on the scientific results rather than the system environments.



NASA Workflow Tool:
Easy Model Operation

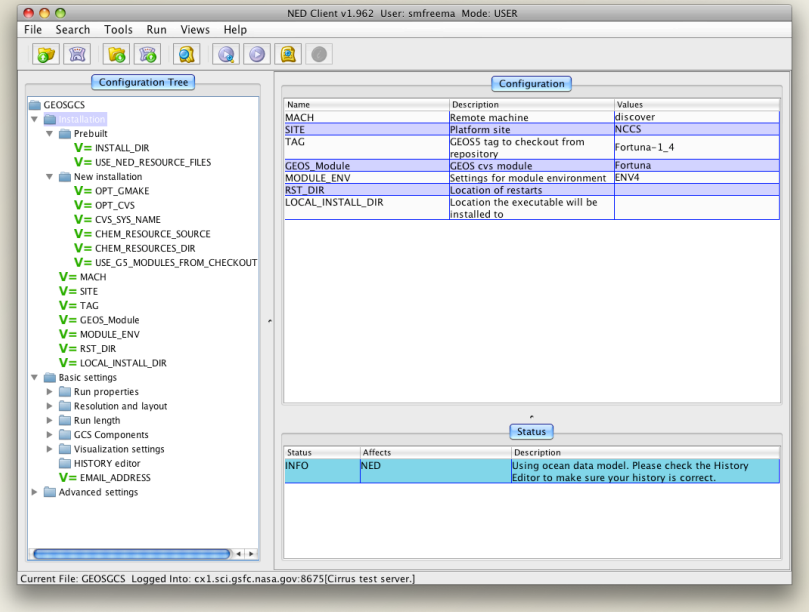
- Configure, run and monitor models from a GUI
- Displays configuration and execution steps
- Removes need for user to have to manage:
  system environment, build settings, config files and other details
- Intended for using models -- not a replacement for highly customized modeling activities

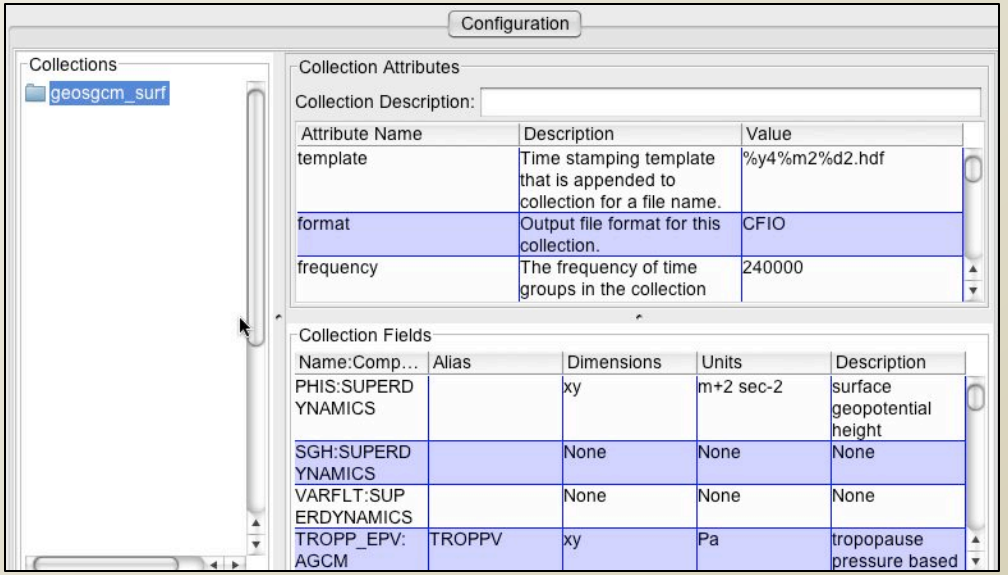




NASA Workflow Tool: Configuration Assistance

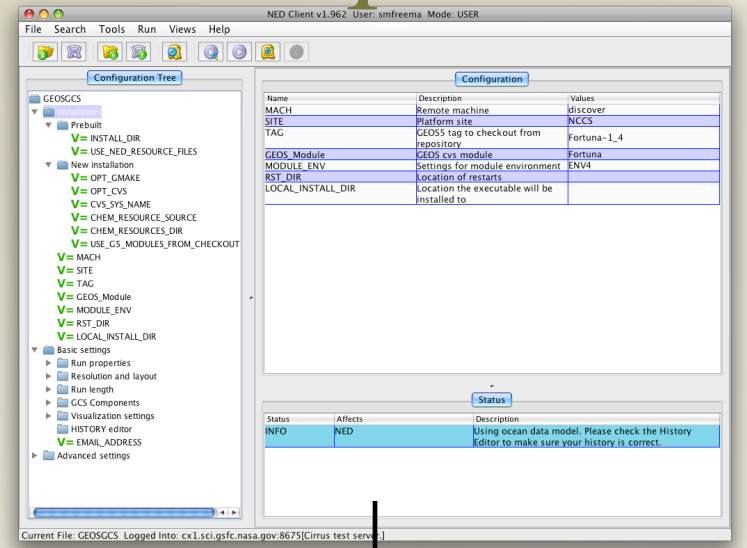
- Model options are organized and presented for configuration and running the model
- Easier for non-specialists to run models
- Can script model configuration actions
  - *Example*: Detect configuration error before running

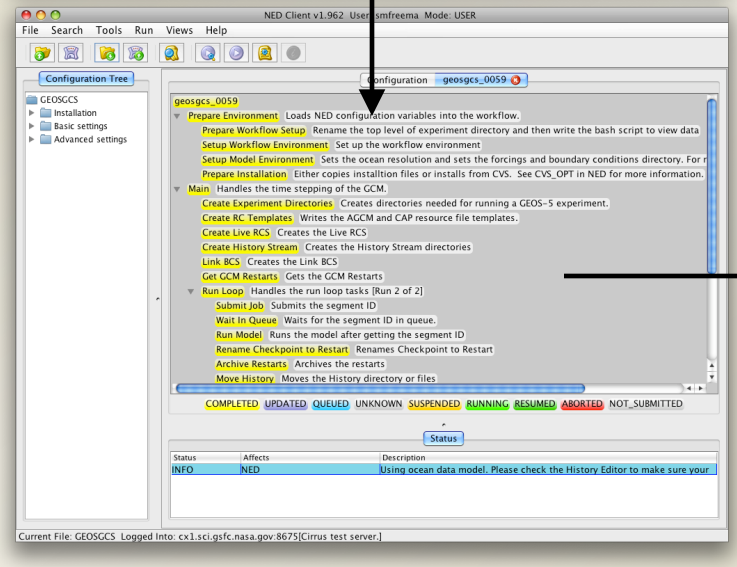


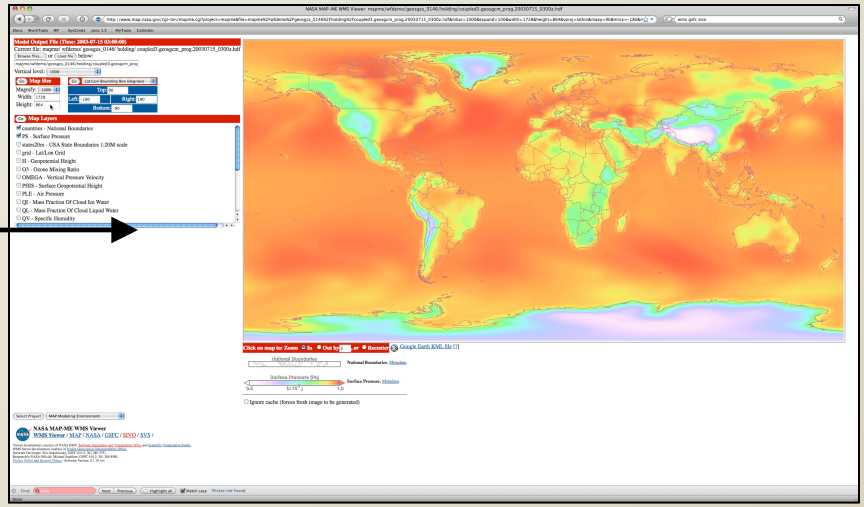


NASA Workflow Tool:

Experiment Management



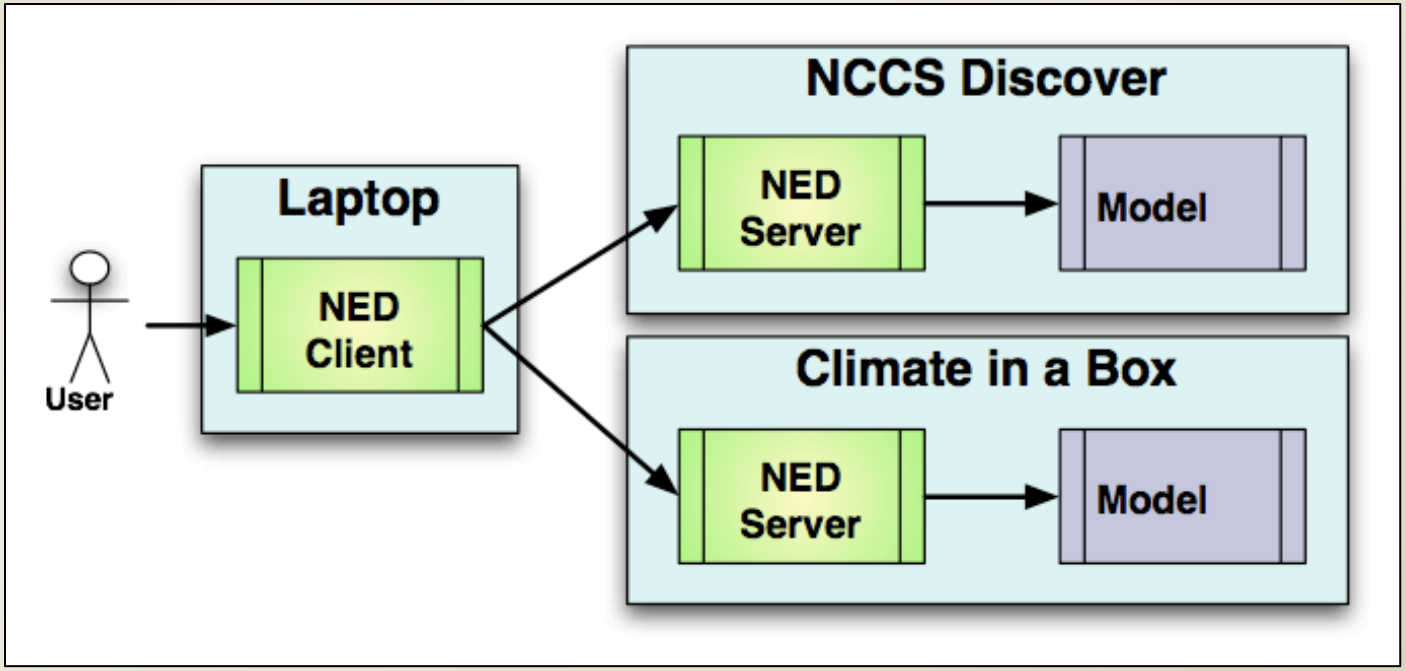




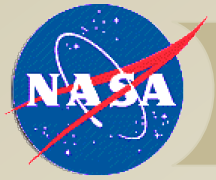
- Save experiment configurations outputs for each run
- Load old experiments and rerun
- Allows experiments to be easily shared and repeated

NASA Workflow Tool: Supports Multiple Sites

- Use the client and switch to a new location
- Use similar configuration settings to re-run the model at the new location



NASA Experiment Designer (NED)

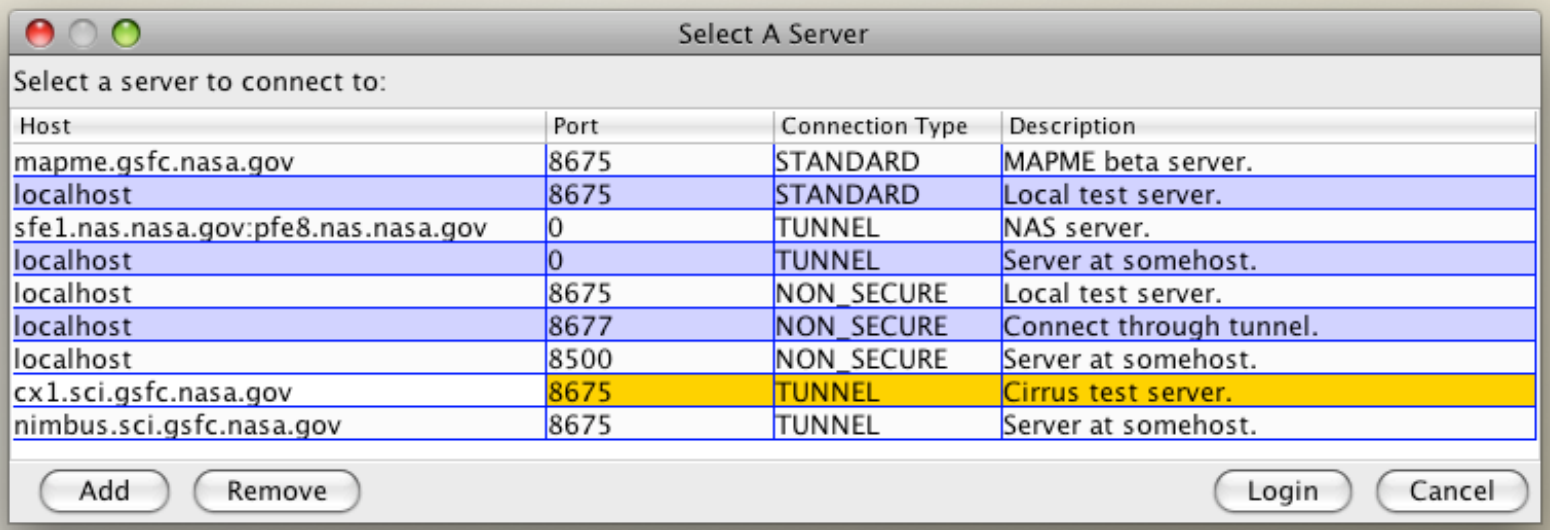


NASA Experiment Designer (NED)

- Consists of a client GUI software and server that runs the workflows underneath.
- Allow user to quickly and easily set up and run an existing workflow
- NED GUI application provides a simple user interface for:
  - Configuring and submitting the workflow
  - Monitoring tasks and task logs as they execute
  - Controlling the tasks in the workflow

NED Server

- The NED client software connects to a NED server on the host machines
  - NED provides a GUI to manage logins
  - Select a site to connect to, such as “mapme”
  - Must authenticate at the particular site



Using the NED Client: Walk-through

- Starting the NED client
- Window layout
- Loading a configuration (local and remote)
- Making changes to a workflow
- Saving a workflow configuration
- Run modes: As Generic or Yourself
- Submitting a workflow
- Monitoring and controlling the workflow

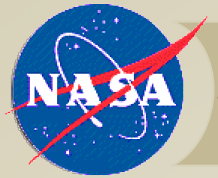
Running a GEOS-5 Workflow

GEOS5 Workflow

- Supports running on both Discover and CIB systems
- Written in Python
- Supports GEOS5 versions on Discover: MERRA/Eros 25, Fortuna 1.x and 2.0
  - Fortuna 2.1 forthcoming
  - Supports checkout-and-build from GEOS5 repository
- Workflow includes basic restarts for a coarse model run

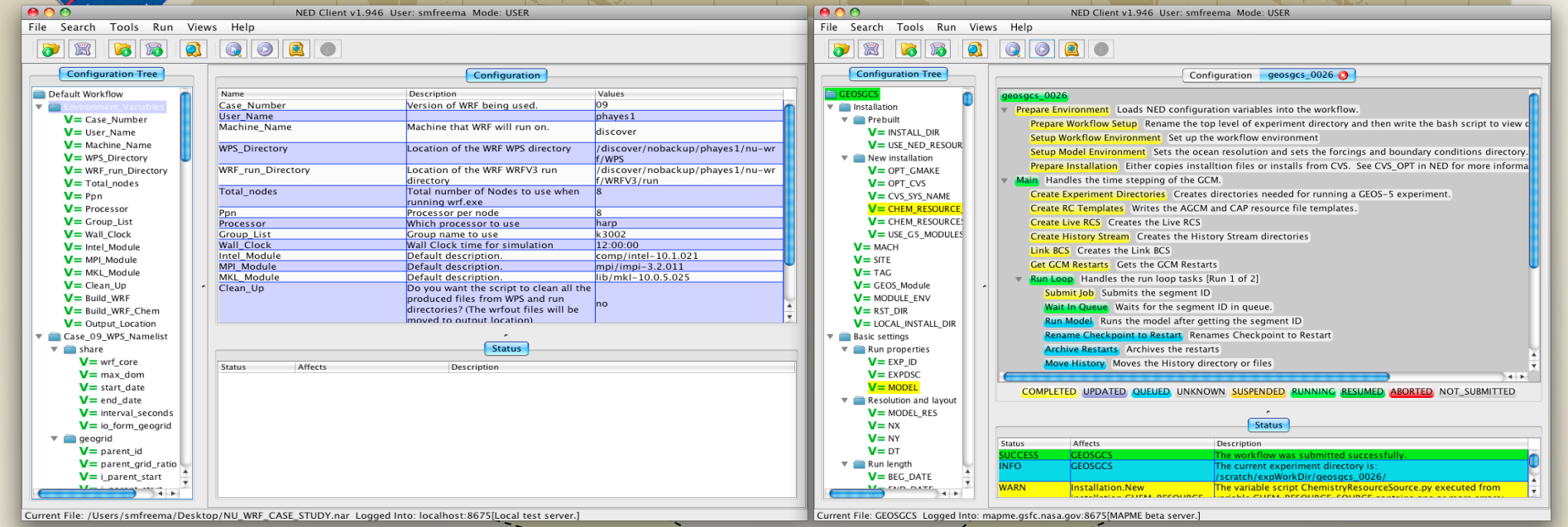
GEOS5 Workflow: Hands-on Demo

- Loading the latest workflow configuration
- GEOS5 settings, History Editor
- Connect/run the model from CIB
  - Accessing the results
- Connect/run the model from Discover
  - Accessing the results
- Visualization
  - Using WMS to view results
  - Forthcoming NED feature: view generated images



Creating a Workflow
Part I: Overview

Workflow Tool Components



File or Searchable DB

Previous Configuration

Version-Control Repository

Architecture/Scripts

NED
(Configure, Monitor, Control)

New Configuration

LWWE
(Execute)

Running Workflow

Parts of a Workflow: Configuration File

- Typically what is displayed to the user when they are setting up a workflow to run
- Contains:
  - Options that a developer wants to make available to users (e.g. climate model options) and how to display them to the user)
  - Metadata about the workflow itself (e.g. workflow name and where it exists)

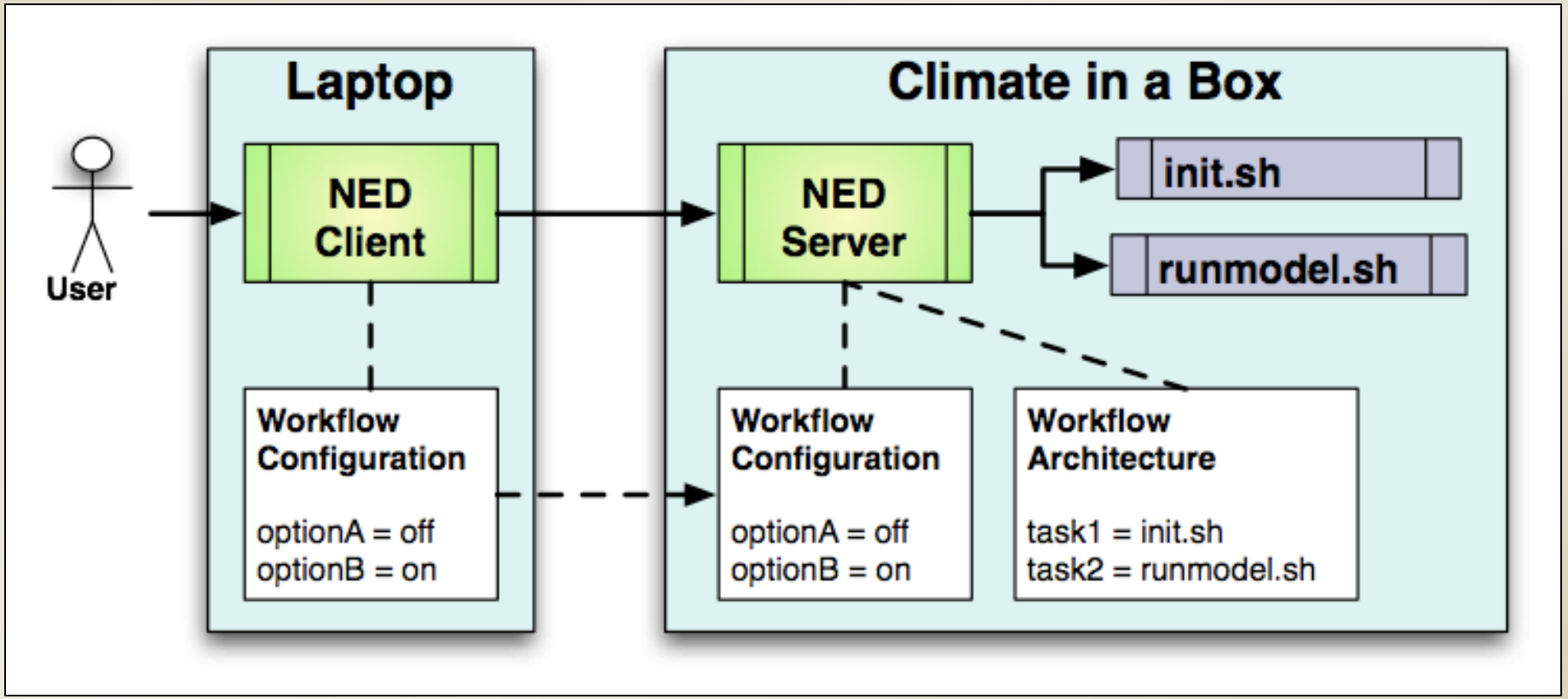
Parts of a Workflow: Architecture File

- Typically displayed to the user during the execution of the workflow (after submission)
- Contains all the task information necessary to run a workflow (e.g. building a code and running it)

Parts of a Workflow: Other Elements

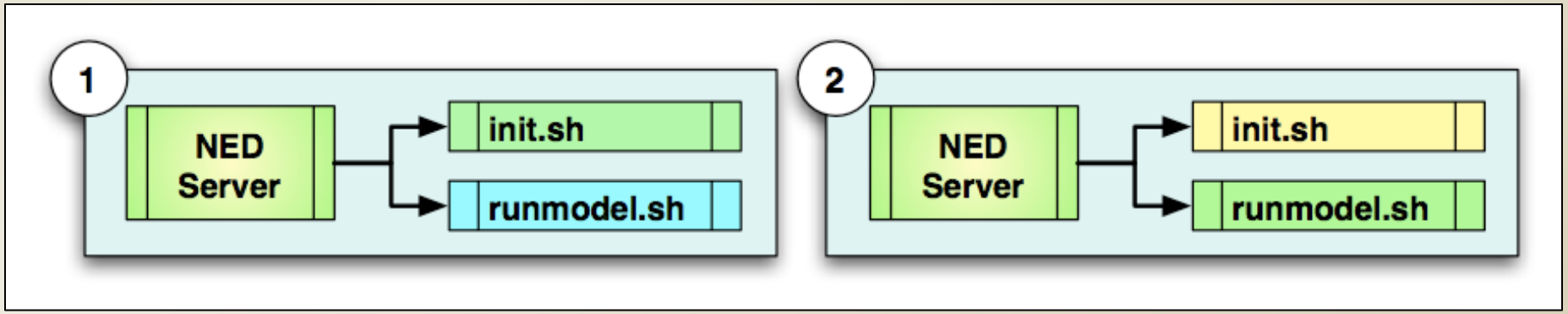
- Workflow tasks do activities such as run system commands, shell scripts, or interact with the workflow
  - These are the low level tasks that perform the work for workflow

Parts of a Workflow: Example



Task Dependencies

- Workflows operate based on task dependencies
- Task dependencies indicate when each operation occurs
- Example:
  - Task “Run Model” depends on “Init” finishing successfully



Designing Workflows

- Designing good workflows is not simple!
  - Must provide flexible, organized configuration for users
  - Must automate the steps a user normally takes
  - Finding problems can be tricky
- The NASA Workflow suite provides tools to help in designing workflows

Features for Workflow Developers

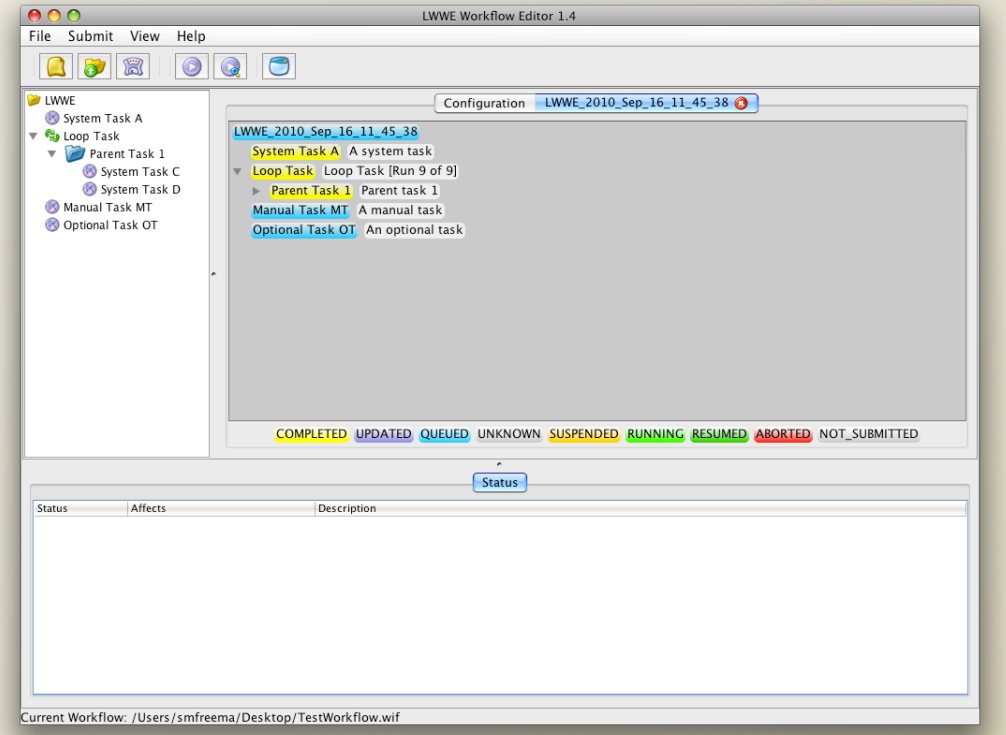
- "Design Mode" feature of the **NASA Experiment Designer** (NED) client software
  - This allows the developer to create a user-friendly configuration
- **Light-Weight Workflow Engine** (LWWE) **Editor**
  - This allows the developer to specify the various tasks that make up a workflow
- **Workflow templates**
  - This allows users to start from a working workflow and expand it into something useful.

NED Designer Interface

- Available through “File / Design Mode” option in NED
- Makes additional options available:
  - Edit overall workflow properties
  - Add, remove and edit workflow variables and groups
  - Develop configuration scripts for advanced configuration options
    - Example: check for incorrect settings

LWWE Editor Tool

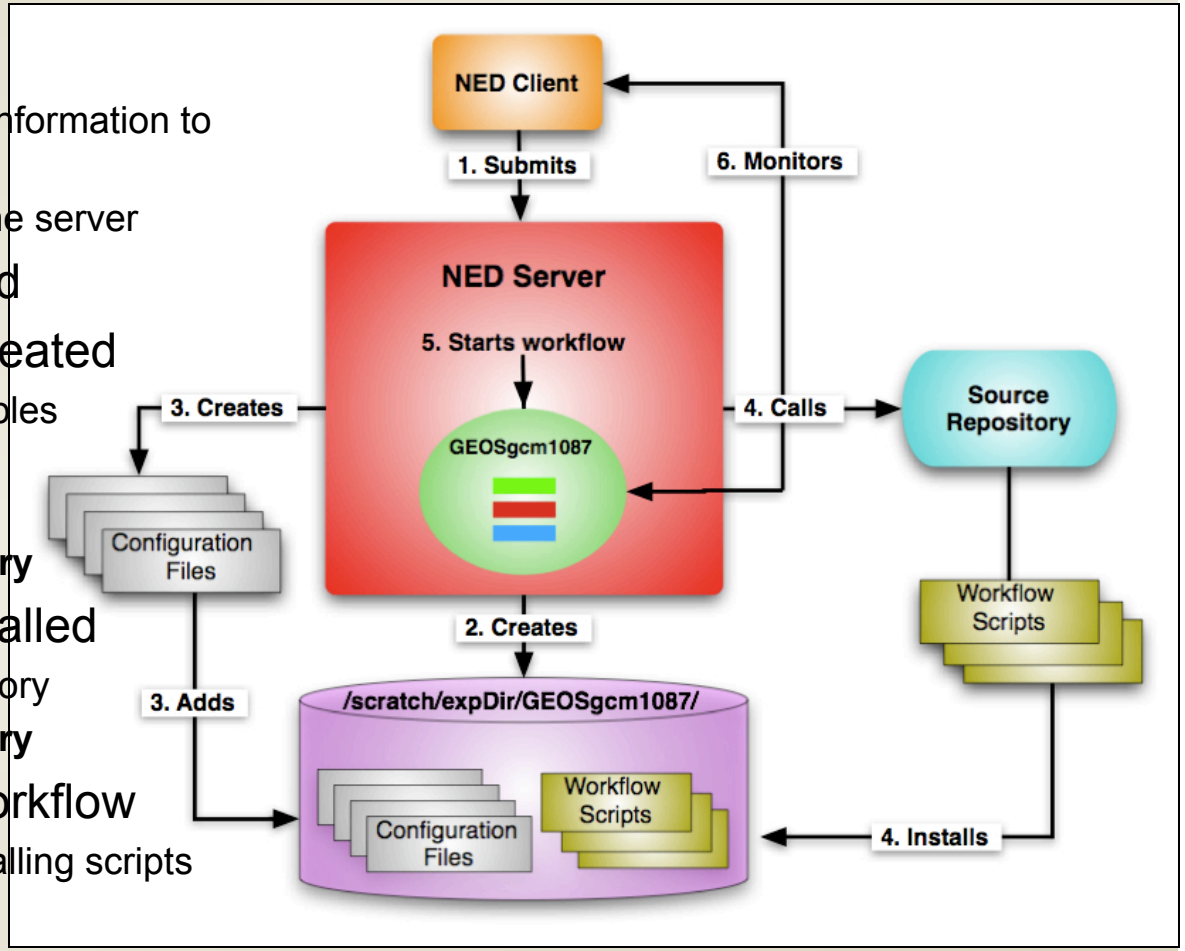
- Allows a workflow developer to quickly construct and validate a workflow architecture
- Test mode allows the developer to test the workflow without running the actual tasks
- Shares a common user interface with NED tool



- WRF workflow demo
- Case Study: WRF model workflow
- Quick demo of LWWE editor

Running a LWWE Workflow in NED

1. User submits workflow
   - NED client sends workflow information to NED server
   - Validation is performed by the server
2. Install directory is created
3. Configuration files are created
   - Define the experiment variables
   - Generated according to the configuration settings
   - Added to the **install directory**
4. Workflow scripts are installed
   - Pulled from a source repository
   - Added to the **install directory**
5. NED server starts the workflow
   - Tasks execute, potentially calling scripts
6. Monitoring occurs
   - Users monitors and controls their workflow



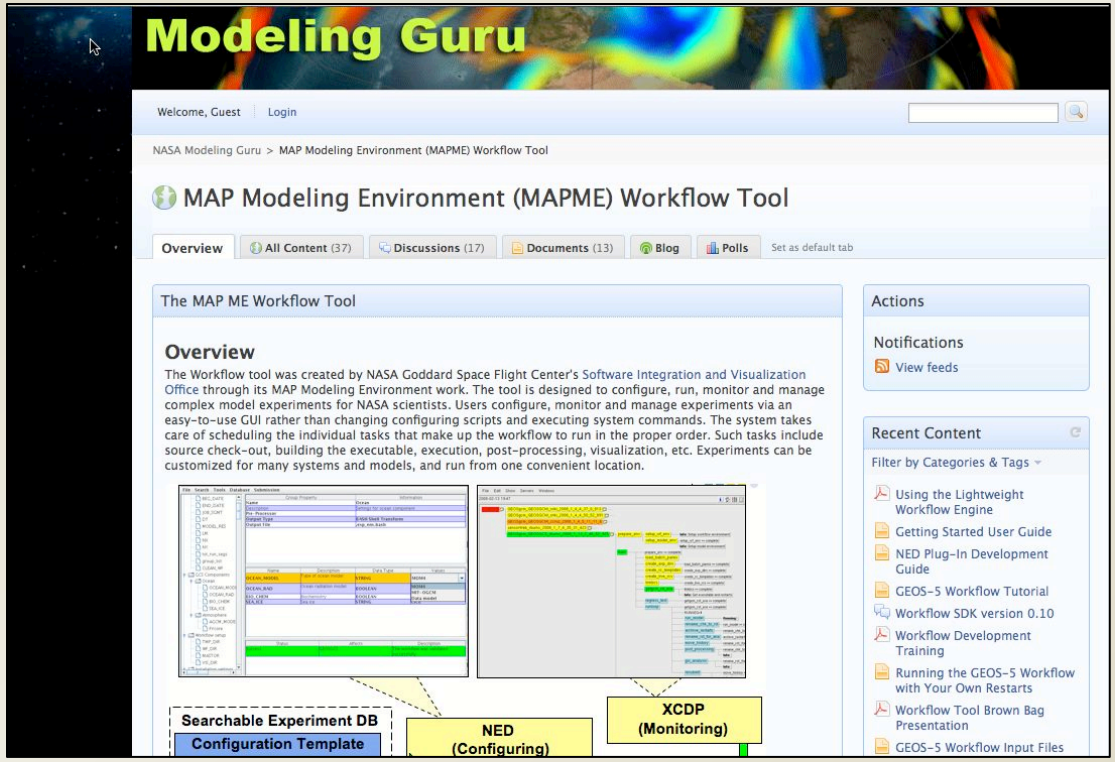
Creating a Workflow
Part II: Hands-on Training

Creating a Workflow Demo

- Demo steps:
  - Access the example workflow
  - Run the example workflow
  - Add a new user setting in the example workflow
  - Add a new task in the example workflow

Workflow References

- **User Guide** and **Developer Guide** available on **Modeling Guru**:



https://modelingguru.nasa.gov/clearspace/community/mapmewkflow